

Un Wiki/IDE pour exploiter le web de données

Pavel Arapov, Michel Buffa et Amel Ben Othmane

Equipe Wimmics, commune aux laboratoires
INRIA et I3S de Sophia Antipolis,
{arapov, buffa, abenothm}@i3s.unice.fr

Résumé : Ce papier décrit WikiNEXT, un moteur de wiki créé pour la rédaction des applications web qui exploitent le web de données directement dans le navigateur web. WikiNEXT est un wiki à la croisée des wikis et des outils de développement en ligne (« web based IDEs¹ »), ce qui fait son originalité. WikiNEXT propose aujourd'hui des fonctionnalités et une interface plutôt orientée vers les développeurs web voulant manipuler des données sémantiques à l'aide des technologies « front end » (JavaScript/HTML5), et offre des moyens pour bénéficier de services proposés par le wiki pour la persistance côté serveur, en fournissant notamment une base de données de graphe compatible RDF/SPARQL 1.1, et une base de données objet.

Mots-clés : wiki, wiki sémantique, IDE basé Web, application Web, Web de données.

1 Introduction

Depuis l'apparition du premier wiki en 1995, créé par Ward Cunningham, (Leuf et Cunningham 2001) de nombreux moteurs de wikis sont apparus, proposant des fonctionnalités communes. MediaWiki est un exemple de moteur de wiki pour faire tourner Wikipedia. Plus tard, les wikis sont devenus très populaires comme des systèmes de gestion de la connaissance. Cette classe de wikis s'appelle les « wikis d'application ». On trouve dans cette catégorie² Confluence, Mindtouch, TWiki. Une autre évolution intéressante des moteurs de wikis, issue du monde de la recherche académique, est apparue en 2005 : les « wikis sémantiques ». Ces wikis sémantiques tels que MediaWiki (Krötzsch et Vrandečić 2006), MoKi (Rospocher et al 2009), OntoWiki (Auer et al. 2006) et IkeWiki (Shaffert 2006), ont permis d'étendre l'approche des wikis classiques avec le contexte sémantique tout en préservant la simplicité et l'essence collaborative des wikis. Récemment, avec le support de la technologie Ajax et le développement de JavaScript, les éditeurs de code source pouvant fonctionner dans une page Web n'ont cessé de s'améliorer. Certains éditeurs comme Cloud9³, Nitrous.IO, jsfiddle.com, et jsbin.com sont très connus par les développeurs.

2 WikiNEXT

WikiNEXT, un mélange entre un IDE basé Web et un wiki classique, permet de développer des applications web qui exploitent le web de données. Plusieurs éditeurs sont fournis : un éditeur WYSIWYG pour des documents classiques, et un éditeur de code pour la partie « application » qui peut être utilisé en mode développement. Le prototype est open source et disponible en ligne⁴. Il inclut aujourd'hui de nombreux exemples et tutoriaux.

¹ IDE = Integrated Development Environment.

² <http://twiki.org>, <https://atlassian.com/fr/software/confluence>, <http://mindtouch.com>

³ <https://c9.io/>

⁴ <http://wikinext.gexsoft.com/>

3 Architecture

La Figure 1 décrit l'architecture logicielle de WikiNEXT. Une page WikiNEXT est considérée comme une application web. Ces applications exploitent des ressources externes (i.e., Freebase.com ou DBPedia.fr) ou internes (données stockés dans la base de données). Une application dans WikiNEXT contient du code HTML/CSS/JavaScript. Dans un paradigme MVC, la partie JavaScript de la page formera « la couche métier et le contrôleur », la partie « vue » se fera dans le code HTML / CSS de la page, et la partie « modèle » sera composée de métadonnées RDF. Chaque page WikiNEXT est associée avec un ensemble de métadonnées qui décrivent ses principales caractéristiques : titre, auteur, contributeurs, dernière date de modification, versions, etc. Mais les pages sont aussi des « containers » : elles contiennent également des métadonnées qui auront pu être rajoutées manuellement ou par une application d'annotation. WikiNEXT s'appuie sur les ontologies disponibles sur schema.org pour décrire la structure des pages, les utilisateurs, les applications web ou le contenu des pages. Les métadonnées générées dans les pages de wiki sont stockées dans un triple store intégré à la partie serveur du Wiki lors de la sauvegarde. Nous utilisons pour gérer les triplets RDF une version améliorée de RDFStore, RDFStore-js, un moteur de graphe compatible SPARQL 1.1, écrit en JavaScript. Le contenu traditionnel de la page est stocké sous forme d'objets dans la base de données MongoDB⁵. Nous l'utilisons également comme couche de persistance pour RDFStore-js.

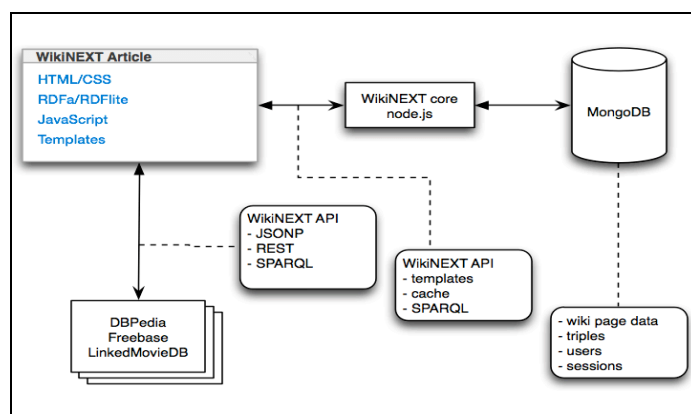


Figure 1: Architecture logicielle de WikiNEXT

4 Démonstration

Dans cette section nous décrivons un scénario d'usage complet : le développement d'une application qui utilise le SPARQL endpoint de DBPedia.org pour récupérer des données RDF concernant des villes (description, population, photos, etc.).

Un formulaire de saisie permet à un utilisateur de saisir le nom des villes qu'il souhaite afficher. Le résultat de la requête est affiché en deux modes : le mode par défaut « sparql On Fly » et en utilisant des pages modèles (templates) basées sur le framework mustache.js (voir la différence entre les deux modes d'affichage dans Figure 2).

Ces résultats sont utilisés pour créer à la volée des pages du wikis –une par ville–, basée sur un modèle de présentation contenant aussi les métadonnées, lui aussi créé dans le wiki comme une page template. Le template gère à la fois l'affichage des données extraites à partir de

⁵ <http://mongodb.org>

DBPedia et les annotations RDFa. Les pages créées sont annotées et les annotations sont sauvegardées dans le triple store RDF du wiki.

Pour démontrer la réutilisation des données dans notre wiki, on a créé une page WikiNEXT qui réalise un « mashup sémantique » construit en requêtant la base de connaissances globale de WikiNEXT sur les villes stockées au lieu de requêter le web de données. Le résultat est affiché dans une carte présentant l'ensemble des villes qui ont été récupérées, avec un résumé et des photos de chacune d'entre elles.

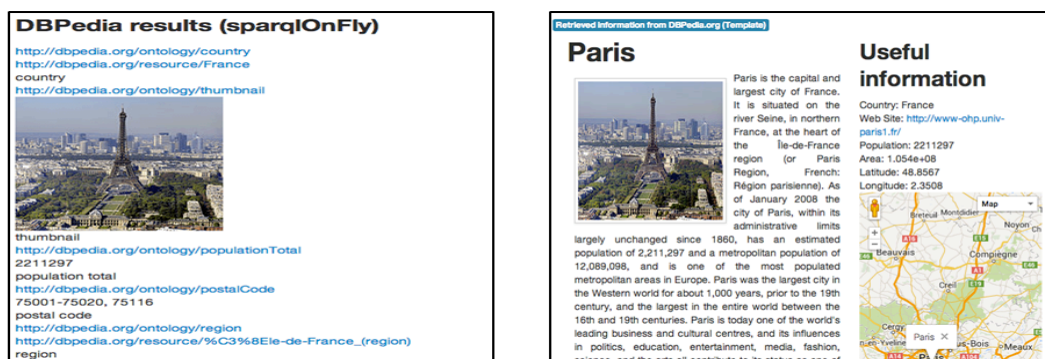


Figure 2: Différence entre l'affichage en mode "sparql On Fly" (à gauche) et en utilisant un template (à droite)

Cette application est disponible en ligne sur le site de WikiNEXT⁶, chaque utilisateur peut consulter le code de l'application, le modifier ou la cloner pour avoir sa propre version.

5 Conclusion

Nous avons présenté WikiNEXT qui est un mélange entre un wiki sémantique et un IDE basé Web pour exploiter les données liées. Nous avons également présenté l'architecture globale de notre application. La démonstration que nous proposons consiste à présenter les principes de WikiNEXT pour la rédaction de documents et des applications exploitant le web de données directement dans le navigateur. Nous vous proposons de reproduire en direct, le scénario de la ville décrit dans ce papier, qui montre comment WikiNEXT simplifie la programmation de ces applications, par rapport à des solutions classiques.

Références

- Auer, S., Dietzold, S., & Riechert, T. (2006). OntoWiki—A tool for social, semantic collaboration. In *The Semantic Web-ISWC 2006* (pp. 736-749). Springer Berlin Heidelberg.
- Krötzsch, M., Vrandečić, D., & Völkel, M. (2006). Semantic mediawiki. In *The Semantic Web-ISWC 2006* (pp. 935-942). Springer Berlin Heidelberg.
- Marco Rospocher, Chiara Ghidini, Viktoria Pammer, Luciano Serafini, Stefanie N. Lindstaedt: MoKi: the Modelling wiKi. *SemWiki 200*
- Schaffert, S. (2006, June). IkeWiki: A semantic wiki for collaborative knowledge management. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE'06. 15th IEEE International Workshops on* (pp. 388-396). IEEE.

⁶ <http://wikinext.gexsoft.com/wiki/519e04c580194c4178000001>